



# Spatio-temporal modelling of *Glossina palpalis gambiensis* and *Glossina tachinoides* apparent densities in fragmented ecosystems of Burkina Faso

Luigi Sedda, Laure Guerrini, Jérémy Bouyer, Naférima Koné and David J. Rogers

L. Sedda (luigi.sedda@zoo.ox.ac.uk) and D. J. Rogers, Spatial Ecology and Epidemiology Group, Dept of Zoology, Univ. of Oxford, Tinbergen Building, South Parks Road, Oxford, OX1 3PS, UK. – L. Guerrini, Centre de Coopération Internationale en Recherche Agronomique pour le Développement, FR-34398 Montpellier Cedex 5, France, and Centre International de Recherche-Développement sur l'Élevage en Zone Subhumide, 01 BP 454 Bobo-Dioulasso, Burkina Faso. – J. Bouyer, Centre de Coopération Internationale en Recherche Agronomique pour le Développement, FR-34398 Montpellier Cedex 5, France, and Centre International de Recherche-Développement sur l'Élevage en Zone Subhumide, 01 BP 454 Bobo-Dioulasso, Burkina Faso. – N. Koné, Centre International de Recherche-Développement sur l'Élevage en Zone Subhumide, 01 BP 454 Bobo-Dioulasso, Burkina Faso.

Tsetse are the insect vectors of the African trypanosomiasis. As with many diseases, transmission of trypanosomiasis varies through space and time. Capturing the variation of both vector and disease has, in the past, been attempted separately in the space and time dimensions, usually using deterministic techniques. Very few efforts have used space-time covariation and have hence missed any correlations that may exist between variation in these two dimensions.

Here we propose two novel approaches to space-time analysis derived from space-time geostatistics in a kriging framework. The approaches were developed through analysis of a dataset recording the Apparent Density of *Glossina palpalis gambiensis* and *Glossina tachinoides* (Diptera: Glossinidae) in three riparian sites in Burkina Faso over 15 months between 2006 and 2007. This site is fragmented due to human activity in the area.

The first approach, Space Time Ordinary Kriging, does not consider the effect of fragmentation. It is used as a benchmark to test the increased explanatory power of the second method, which does account for fragmentation. The second method, Regression Space Time Simple Kriging, is a distinct improvement over the first approach because it allows for a spatial trend in the mean trap catch; this trend is related to, and later predicted from, environmental co-variables.

The results indicate the presence of space and time effects on tsetse distribution, dependent on the size of the habitat fragmentation patches. These effects occur at relatively small geographic scales within a season. Whilst such variation has long been suspected, the new methods presented here are able to quantify this variation precisely, so that seasonal and spatial comparisons can now be made both within and between species.

The tsetse transmitted trypanosomiasis, causing “nagana” in domestic animals and sleeping sickness in humans, have had an enormous social and economical impact on development in Africa (Maudlin 2006). Tsetse belong to the genus *Glossina* (order: Diptera, family: Glossinidae) which is divided into three groups: *Fusca*, *Morsitans* and *Palpalis* (Krafsur 2009). The latter contains *Glossina palpalis*, a species mainly confined to the rivers of West Africa, and *G. tachinoides* with a somewhat wider geographical range (Rogers and Robinson 2004). Tsetse flies of the *palpalis* group take blood meals from a wide variety of animals, including reptiles and, as a result, are slightly less susceptible than are *G. tachinoides* (which prefers to feed on mammals) to habitat modification (Hendrickx et al. 1999), as long as sufficient shade remains for shelter and for the

deposition of their larvae (Jordan 1986). The abundance of these and other tsetse flies thus depends on the presence of hosts and the existence of suitable habitat, which together influence the spatial and temporal dynamics of the species.

Considering first only the spatial dimension, the prediction of a species' distribution over very large areas (e.g. continents or countries) may be made using deterministic techniques where the effects of local autocorrelation are absorbed by spatial trends (Rogers et al. 1996, Hendrickx et al. 1999, 2001, Odit et al. 2006, Guerrini and Bouyer 2007). At a finer spatial scale, and when temporal trends are also important, this approach is no longer adequate (Wagner and Fortin 2005). To understand tsetse dynamics properly at these finer scales we first need to understand the spatial and temporal relationships between flies and

their environment. In the past this understanding has come from sampling flies at fixed points (traps) or along regularly worked transects in what is considered to be suitable tsetse habitat. A trap catch, or transect count, is a relative, not absolute, measure of fly abundance that is probably biased seasonally. Such catches are referred to in the tsetse literature as “Apparent Densities” ( $AD$ ).

One of the first stochastic approaches to tsetse modelling was applied by Yu et al. (1995) in an epidemiological space-time model based on the long accepted assumption that fly movement is random in homogeneous areas (Hargrove 2000, Bouyer et al. 2007a). Spatio-temporal autocorrelation was not taken into account. Recently, Odulaja et al. (2001), Sciarretta et al. (2005) and Cano et al. (2007) studied the spatial and temporal variability of  $AD$  considering the two dimensions separately, but calculating autocorrelation in the space domain only. Even if these studies did not consider the covariance function in both space and time, they demonstrated the advantage of using a geostatistical approach, at least for the space dimension, as can be found for other fly genera (i.e. Hwang and Turner 2005) and in general for insect population distributions (Arbogast et al. 1998).

de la Rocque et al. (2001) analysed the influence of landscape disturbance on the distribution of *palpalis* group tsetse flies over 15 yr in Burkina Faso and found that whilst overall tsetse abundance was unchanged their spatial distribution was altered. Tsetse populations decreased in some areas, owing to vegetation clearing and farming system modification, and increased in others, particularly in those areas with a continuous presence of cattle, on which the flies fed.

The present study was carried out in an area very similar to de la Rocque’s, along the Mouhoun River in Burkina Faso. Here *G. palpalis gambiensis* and *G. tachinoides* are mostly restricted to riparian vegetation, which has been more or less modified by humans (Van den Bossche 2001). Elsewhere in Africa, close (peri-domestic) association of flies with humans and their cattle has been shown locally to increase infection rates (Rawlings et al. 1993), and this has also been observed along the Mouhoun River (Guerrini and Bouyer 2007).

The question we answer in this paper is whether there is a connection between the variation in space and variation in time of the abundance of tsetse species living in human-fragmented habitat. Clearly, if there is a connection it indicates both that tsetse are responsive to such fragmentation in rather subtle ways and that such responses should be considered in determining tsetse monitoring and control strategies in the future. The larger the area over which space-time connections are demonstrated naturally, the larger must be the area considered for tsetse control by whatever means available.

We present two space-time geostatistical kriging methods (Christakos 2000, Gneiting 2002, De Iaco et al. 2005), the second derived from, and an extension of, the first (Fig. 1). Kriging is a standard spatial interpolation procedure used to predict the value of a variable at a certain location through the use of a system of linear equations in which the weights are determined by the covariance function (Cressie 1993). The first method presented here is called standard space-time ordinary

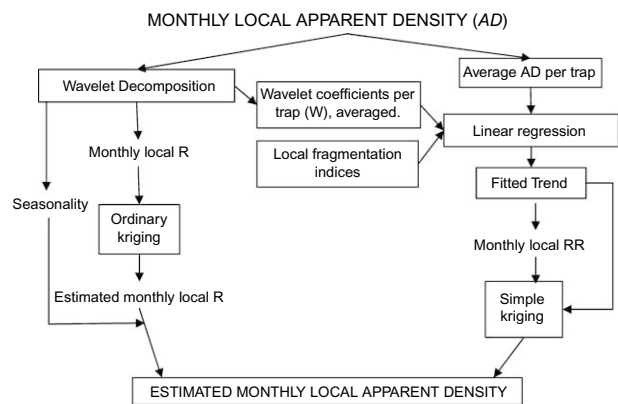


Figure 1. Model frameworks applied: STOK (on the left) and RSTSK (on the right).

kriging (STOK) and is applied to the residuals ( $R$ ) produced by removing seasonality from tsetse catches through a wavelet technique described below.

The second method allows mean tsetse abundance to vary through space. The only estimate we have of this variation is derived from the sample data, which are first modelled with regression methods to produce an estimate of the trend. Removing this trend from  $AD$  produces the residuals ( $R_R$ ) used to model the space-time covariance function, which is then applied to predict residual abundance through space and time. This technique is called regression space-time simple kriging (RSTSK).

In ordinary kriging the mean is unknown but assumed constant (i.e. mathematically stationary) while in simple kriging the mean is known and may vary through space (in a way that is estimated from the data). In this paper the term “seasonality” refers to the temporal (seasonal) cycle of  $AD$  variation, and the term “trend” refers to local variation of the mean through space. We use the phrase “experimental covariance” to refer to the covariance derived directly from the data, and “theoretical covariance” to refer to the model chosen to fit the experimental covariance. Ordinary and simple kriging use a theoretical covariance model defined as “nonseparable” because not separable in space and time (Cressie and Huang 1999).

The two techniques differ importantly in the approach to deriving the residuals ( $R$  or  $R_R$ ), to which the analyses are applied. Since the STOK approach assumes a constant mean, de-trending involves only de-seasonalising the data in order to stabilise the variance. In contrast, RSTSK uses linear regression (Minasny and McBratney 2007) to de-trend  $AD$  where the predictors are a measure of seasonality (see below) and the (locally variable) fragmentation index. We opt to decompose the  $AD$  into its components (seasonality, trend, stochastic process) in order to be able to analyse each one separately.

Finally, the analysis compares and tests the performance of the two procedures using leave-one-out cross validation, the correlation coefficient, and the  $D^2$  statistic (Bogaert 1996). The main abbreviations used in the text are listed in Table 1.

Table 1. Common abbreviations used in the text.

$\alpha$	Smoothing parameter for the time domain. Used in the autocovariance function.
$\beta$	Dimensional parameter. Used in the autocovariance function.
$c$	Scaling parameter for the space domain. Used in the autocovariance function.
$s$	Site index, $s = 1$ to 61.
$t$	Month index, $t = 1$ to 15.
$h$	Spatial lag (Euclidean distance between two sites), metres.
$u$	Temporal lag (absolute difference in months between two sites), month.
$AD$	Apparent Density of tsetse, a standard measure of the relative (not absolute) abundance of tsetse flies usually (though not always) caught by traps.
$CA$	Class Area. It is the total area of suitable habitat, $m^2$ .
$ME$	Mean Error. Used in cross validation.
$MSE$	Mean Squared Error. Used in cross validation.
$MSDR$	Mean Squared Deviation Ratio. Used in cross validation.
$MPS$	Mean Patch Size of forest, $m^2$ .
$MODWT$	Maximum Overlap Discrete Wavelet Transform.
$PSSD$	Patch Size forest Standard Deviation.
$R$	Residuals from removing the seasonality in $AD$ . Used in STOK.
$R_R$	Residuals from removing the spatial trend in $AD$ . Used in RSTOK.
$RSTSK$	Regression Space time Simple Kriging.
$STOK$	Space time Ordinary Kriging.
$T$	Total length of the time series (15 months).
$W$	Wavelet coefficients.

## Materials, methods and preliminary analysis

### Study area and tsetse sampling

Tsetse flies were sampled in three different areas: Douroula, Kadomba and Boromo along the Mouhoun River (Fig. 2), between the three major urban areas of Bobo Dioulasso (south-west), Dédougou (north) and Boromo (south-east), ca 100 km apart from each other. The sampled areas are all in the Sudan vegetation zone of bush savannas, with forest stands along the watercourses. The only tsetse species found are *G. tachinoides* and *G. p. gambiensis*.

Flies were caught three times per month from May 2006 to July 2007 using 61 fixed georeferenced unbaited biconical traps with ca 20 traps per site, placed at distances



Figure 2. The study area and sampling locations (Douroula, Boromo and Kadomba) in Burkina Faso, West Africa.

apart of between 100 and 200 m; a suggested distance for efficient trap deployment in tsetse monitoring (Kuzoe and Schofield 2005). This sampling design is unlikely to be able to control the tsetse population because it covers a small portion of the river course only.

There was considerable variation in the abundance of the tsetse species in the three study areas. For example, 5000 *G. p. gambiensis* were trapped in Kadomba but none at all in Boromo. *Glossina tachinoides* was very rare in both of these sites, but much commoner than *G. palpalis* in Douroula. Given this variation, the analyses were carried out only for *G. p. gambiensis* in Douroula and Kadomba and *G. tachinoides* in Douroula, there being too few data for the other sites.  $AD$  for each species in each area was calculated as the total count of flies (male plus female) captured by each trap per month. Male and female numbers were merged because of the very small sample size in many trap sites; in fact, for each sex on its own 44 and 88% of trap catches were zero across all sampling occasions. This result reflects the low overall abundance of flies in this area and also trap efficiency (the traps were unbaited).

Trap catching can itself have a direct effect on the population density of tsetse (e.g. in areas with few flies, or areas that are sampled very frequently and intensively) thus complicating statistical analyses (Brightwell et al. 1997). The trap effect was tested at each location,  $s$ , using the average of the differences between successive departures,  $\gamma$ , of the actual  $AD$  for each trap sample  $s$ , taken at time  $i$ , from the average  $AD$  for that trap calculated for a short, fixed period,  $T_0$ , of the first three months of trapping (corresponding to the conclusion of the rainy season):

$$\gamma(i; s) = \left[ \frac{1}{T_0} \sum_{t=1}^{T_0} AD_s(t_0) \right] - (AD_s(i))$$

$$\text{for } i = (T_0 + 1), \dots, T \quad (1)$$

$$\gamma(i) = \frac{1}{S} \sum_{s=1}^S \gamma(i; s)$$

where  $T$  is the total time length (in month) and  $S$  is the total number of traps for each area.

Given the results shown in Fig. 3, no influence of the traps on tsetse abundance is detectable (i.e. no negative trend is present), a result that can be accounted for by the low trap density.

### Fragmentation classification

A 200 m wide buffer around the river-course, cut in sections of ca 500 m length, was generated to produce habitat fragmentation indices at each trap site. The grid used in previous research (Guerrini et al. 2008) subdivides riverine forest along the Mouhoun River into 6 fragmentation classes. These classes were identified using a Partitioning Around Medoids approach (Kaufman and Rousseeuw 1990) applied to the forest vegetation classification extracted from supervised classification of Landsat 7 ETM+ imagery (for more details see Guerrini et al. 2008). The following indices were computed for each forest vegetation type: 1) the class area ( $CA$ ), corresponding to the total area

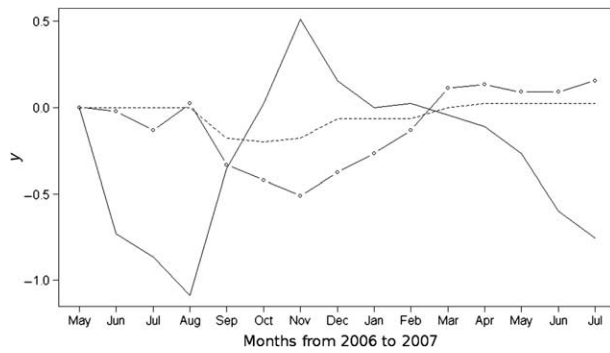


Figure 3. Trap effect analysis from May 2006 to July 2007.  $\gamma$  is the average departures of  $AD$  from the mean of the first three months. Here three examples are presented: *G. palpalis gambiensis* in Kadomba (solid line) and Douroula (dashed line); *G. tachinoides* in Douroula (points and lines).

of suitable habitat; 2) the number of patches of suitable habitat (Huang et al. 2006); 3) the mean patch size ( $MPS$ ); 4) the patch size standard deviation ( $PSSD$ ).

The first index is not a measure of fragmentation but a measure of the total amount of suitable tsetse habitat which has in the past been correlated with tsetse abundance in Nigeria (Bourn 1983). The second index is a common fragmentation metric. Its correlation with  $AD$  in the present study area was very low (Table 2), probably because there was not much variation in this parameter within each area. The variation in size within the single patches seems to drive  $AD$  fluctuations.  $CA$ ,  $MPS$  and  $PSSD$  were significantly linearly correlated with  $AD$  and were therefore used in the RSTSK method as covariates in the linear regression to predict local mean  $AD$ .  $CA$  shows the highest correlation with  $AD$ , with similar values for both species in Douroula. In Kadomba (where only *G. p. gambiensis* was caught) the correlations are very low.

## The general model and its components

This section defines the overall approach to modelling the  $AD$  data, which is shown graphically in Fig. 1. Subsequent sections give more detail of each of the components introduced here. The first model step for each method was to remove seasonality from the data by applying a wavelet decomposition method to the  $AD$  for each trap. Wavelets are filters that decompose a signal into trends and residuals at different scales (temporal resolutions in this case) (Debnath 2002). Certain wavelet components describe the seasonality of the data, the extent of which is captured by the wavelet coefficients ( $W$ ).

Two sets of residuals were derived from the wavelet fits to the trap data. One set ( $R$ ) was derived assuming a stationary spatial mean (i.e. across all traps at any one time). The other set ( $R_R$ ) was derived allowing the mean trap catch to vary spatially in a way described by a linear regression model using various fragmentation indices and mean wavelet coefficients as independent variables. Clearly one would expect smaller residuals in the latter case, since the regression model allows for non stationarity of the mean catch.

The second model step was then to fit the experimental covariance of these residuals ( $R$  or  $R_R$ ) using an appropriate theoretical covariance function.

The third step of the model was to predict the variation in  $AD$  across space and through time using the results of the two kriging approaches (ordinary kriging for  $R$  and simple kriging for  $R_R$ ). Predictions using the STOK approach were made by combining the wavelet-predicted seasonality of  $AD$  with the geospatial variation in  $AD$  predicted from the ordinary kriging results. Predictions from the RSTSK approach combined the regression-predicted trend in mean  $AD$  with the geospatial variation in  $AD$  predicted from the simple kriging results.

Analysis was performed using the software R-2.6.1 (R Development Core Team 2008), with the package waveslim (Whitcher 2007) for wavelet analysis and with modified routines from the RandomFields package (Schlather 2001) for covariance testing, kriging and variography.

## Data pre-processing and filtering

Tsetse Apparent Density is strongly affected by the seasonal periodicity of the habitat (wet, dry, hot and cold seasons), so that monthly  $AD$  data have both high Skew and Kurtosis (Table 3). Figure 4 (top) shows how the  $AD$  of each species varies seasonally but in rather different ways at the different capture sites, possibly as a result of seasonal fly movements from one site to another. The strong variation in  $AD$  between sites (Fig. 4, upper for individual values and lower for the mean value) required the removal of seasonality from each trap time series independently before further analysis.

Two methods are commonly used to remove seasonality from time series: a sine/cosine method and wavelet method. The latter was chosen because the sine/cosine method (Diggle 1990) produced a worse fit of the data (Table 3).

## The wavelet transform method

Wavelet methods offer more flexibility in fitting data than does the sine/cosine method, because the high and

Table 2. Correlation coefficients between the global mean Apparent Density across 15 months of the two fly species and the four fragmentation indices ( $CA$ , class area;  $MPS$ , mean patch size;  $PSSD$ , patch size standard deviation;  $NuP$ , number of patches of suitable habitat).

Species	Site	$CA$	$MPS$	$PSSD$	$NuP$
<i>G. p. gambiensis</i>	All	-0.58	-0.38	-0.39	-0.27
	Douroula	0.65	0.61	0.64	-0.19
	Kadomba	-0.02	-0.10	-0.09	-0.14
<i>G. tachinoides</i>	Douroula	0.53	0.50	0.50	-0.25

Table 3. Summary statistics for species and for seasonality analyses that allow the de-seasonalising the data.  $R$  are the residuals produced after MODWT decomposition or fitting a sine-cosine function to the raw data (individual trap Apparent Density per month).  $R_R$  are the residuals calculated after applying a linear regression in RSTSK procedure (see text).  $\mu$  is the mean,  $\sigma^2$  is the variance, SD the standard deviation, Skew the skewness. The following statistics are referred for *G. p. gambiensis* in Douroula and Kadomba, and *G. tachinoides* in Douroula.

Glossina ssp.	Function/data	$\mu$	$\sigma^2$	SD	Skew	Kurtosis
<i>p. gambiensis</i>	MODWT (R)	0.00	227.15	15.07	2.33	15.30
	Sine-cosine (R)	0.89	156.69	12.52	2.72	25.29
	Apparent Density	8.63	332.60	18.23	9.65	70.73
	$R_R$ residuals	0.00	1.88	1.29	0.93	2.43
<i>tachinoides</i>	MODWT (R)	0.00	12.86	3.59	2.24	12.75
	Sine-cosine (R)	-0.20	11.47	3.39	2.96	13.70
	Apparent Density	2.32	11.90	3.44	3.84	24.40
	$R_R$ residuals	0.00	2.74	1.65	0.84	1.72

low frequency components of any time series cycle are resolved separately and adaptively, thus allowing the capture of changes at any possible location. The maximum overlap discrete wavelet transform (MODWT) (Percival and Walden 2000), a non-orthogonal procedure that gives translation-invariance, was used here to determine time-dependent wavelet and scaling coefficients at each point in time for each scale analysed. The values of  $W$  for each trap time series therefore capture the unique seasonal signal at that trap site. Differences in seasonality between trap sites were analysed using an ANOVA test on the MODWT signal. For *G. p. gambiensis* the seasonality

signal is statistically significantly heterogeneous from one trap to another within the Douroula ( $p < 0.001$ ) and Kadomba ( $p < 0.05$ ) traps. These dissimilarities can be interpreted as the different response of this species of tsetse to a defined landscape according to the season. *Glossina p. gambiensis* changes resting places in relation to the season, and colonizes the savannahs neighbouring the river course during the rainy season (Bouyer et al. 2007b). No significance was detected for the MODWT signals for *G. tachinoides* in Douroula ( $p > 0.05$ ). This tsetse species is more resistant to desiccation and is therefore less likely to change habitats with season.

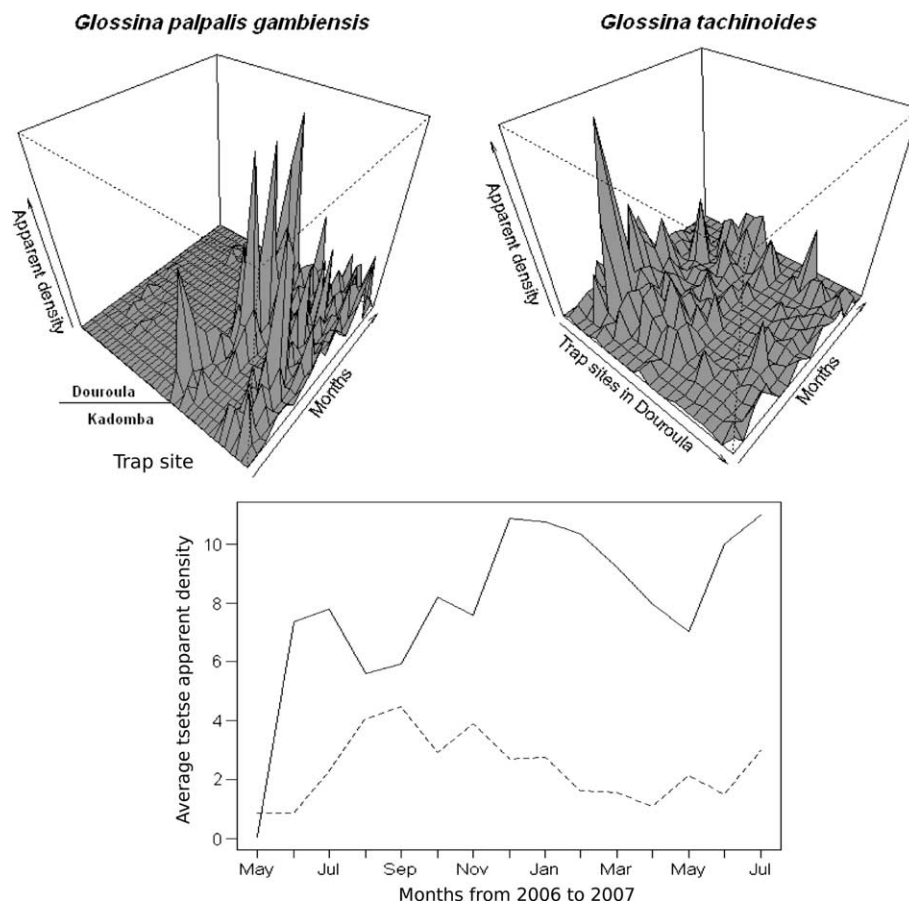


Figure 4. Counts of flies for each trap and month (upper) and mean monthly AD for all the sites combined (lower) from May 2006 to July 2007 (*G. palpalis gambiensis*, solid line; *G. tachinoides*, dashed line).

Residuals ( $R$ ) of the raw data from seasonal estimates derived from the fitted seasonal wavelet components were the inputs into the ordinary kriging (STOK) approach. For the second type of approach developed here (RSTSK), the mean trap  $AD$  were first fitted by linear regression using as predictor variables the mean wavelet coefficients per trap site and three of the four corresponding (i.e. site specific) fragmentation indices ( $CA$ ,  $MPS$  and  $PSSD$ ). The resulting regression was then used to predict the spatial trend of the mean, and the residuals from this trend ( $R_R$ ) were used in the simple kriging approach. The wavelet coefficients and the fragmentation indices used in the regression clearly capture information about the local variation in mean  $AD$  per trap across the sampled areas, and the former also reflect seasonal variation in  $AD$ . The residuals  $R_R$  used in the simple kriging approach are thus less affected by spatial and temporal variation in  $AD$  (since this variation was extracted as the trend of the fitted regression model) than are the residuals  $R$  used in the ordinary kriging approach. Kriging, which assumes a constant spatio-temporal co-variance of the kriged data, should therefore produce better results with  $R_R$  than with  $R$ .

The summary statistics of the residuals produced after fitting a sine-cosine function compared with those obtained from MODWT indicates that the latter method is less biased (lower mean, and less skew and kurtosis) (Table 3), and hence is used throughout here.

### Residual modelling: nonseparable space-time model

The core of STOK and RSTSK methods is the space-time covariance function which defines the relation between the temporal and spatial autocorrelations that are used during kriging to make estimates of the values of the measured variables at points other than those sampled.

In order to apply kriging procedures it was assumed that the variable of interest ( $R$  or  $R_R$ ) at each point is a random variable that can take any value from a distribution of values; i.e. that the data are represented by a space-time random field (Christakos 2000). This method makes two assumptions, of stationarity and isotropy. Stationarity means that the mean and covariances of  $R$  and  $R_R$  are invariant for all translations in space and time. Isotropy means that the covariance function depends only on distance (lag) and not on direction. In the absence of a trend (or after removing it), the first condition is usually met, the second less commonly so. Bearing in mind the different derivations of the residuals  $R$  and  $R_R$  it seems that the latter are more likely to conform to the kriging assumptions.

Putting together these two assumptions, we define the space-time covariance functions,  $\phi$ , as:

$$\phi(h; u) = \frac{\sigma^2}{\psi(|u|^2)^{\frac{d}{2}}} \varphi\left(\frac{\|h\|^2}{\psi(|u|^2)}\right) \quad (2)$$

where  $u$  is the time lag ( $||$  means difference between two times) and  $h$  the space lag ( $||\cdot||$  means Euclidean distance),  $\sigma^2$  is the variance of the process (usually referred to as the sill) and  $d$  is the half number of dimensions (in this analysis  $d$  is always equal to 1.5 because there are two

spatial and one temporal dimensions);  $\varphi(\Delta)$  for  $\Delta \geq 0$  is a monotonic function and  $\psi(\Lambda)$  for  $\Lambda \geq 0$  is a positive function with a completely monotonic derivative.  $\Delta$  and  $\Lambda$  are generic arguments of the functions in which the spatial and temporal ranges are related in some way (specified in more detail below). Technically,  $\phi$ , is a space-time second-order stationary random field (the covariance depends exclusively on the space-time lag) as described by Cressie and Huang (1999) and Gneiting (2002).

Equation (2) is a valid space-time covariance function only if  $\sigma^2 > 0$ . The final fit of this model to the data ( $R$  and  $R_R$ ) depends crucially upon choosing suitably flexible, but also biologically realistic, forms of  $\varphi(\Delta)$  and  $\psi(\Lambda)$ .

In fitting eq. (2) to the data, the following functional form of the space covariance,  $\varphi(\Delta)$ , was used:

$$\varphi(\Delta) = (2^{v-1} \Gamma(v))^{-1} (\Delta)^v K_v(\Delta) \quad (3)$$

Similarly for the time covariance,  $\psi(\Lambda)$ , the following was used:

$$\psi(\Lambda) = (\Lambda^\alpha + I)^\beta \quad (4)$$

where  $\alpha$  and  $\beta$  are smoothness parameters of time, and  $K_v$  is a modified Bessel function of the second kind, of order  $v$ .

Combining eq. (3) and eq. (4) with eq. (2) produces the following explicit parametric function:

$$\phi(h; u) = -\frac{\sigma^2}{(|u|^\alpha + I)^{\beta d}} \frac{1}{2^{v-1} \Gamma(v)} \left( \frac{c \|h\|}{(|u|^\alpha + I)^\beta} \right)^v K_v \left( -\frac{c \|h\|}{(|u|^\alpha + I)^\beta} \right) \quad (5)$$

with  $\frac{c \|h\|}{(|u|^\alpha + I)^\beta} = \Delta$  and  $|u| = \Lambda$

where  $c$  is a non-negative scaling parameter of space. In order to obtain an exponential function (the form of the covariance function that best represents the experimental covariance), a value of  $v=0.5$  was chosen. The above equation then becomes:

$$\phi(h; u) = -\frac{\sigma^2}{(|u|^\alpha + I)^{\beta d}} \exp \left( -\frac{c \|h\|}{(|u|^\alpha + I)^\beta} \right) \quad (6)$$

The model is isotropic in space and time and has four constraints, one for each parameter, that ensures the validity of the covariance function, i.e. that  $\sigma^2 > 0$  (Schlather 2001):

$$\begin{aligned} c &> 0 \\ 0 &< \beta < 1 \\ 0 &< \alpha \leq 2 \\ d &> 1 \end{aligned}$$

Figure 5 shows how the parameters affect the covariance function for 20 spatial lags of 1 unit ( $h$  axis) and 15 temporal lags of 1 unit ( $u$  axis). The parameters modify the surface in a complex way, but using the model in Fig. 5a (with reasonable but arbitrary parameters) as a reference it is possible to compare the major effects. A reduction in the smoothness parameter of time,  $\alpha$ , results in more

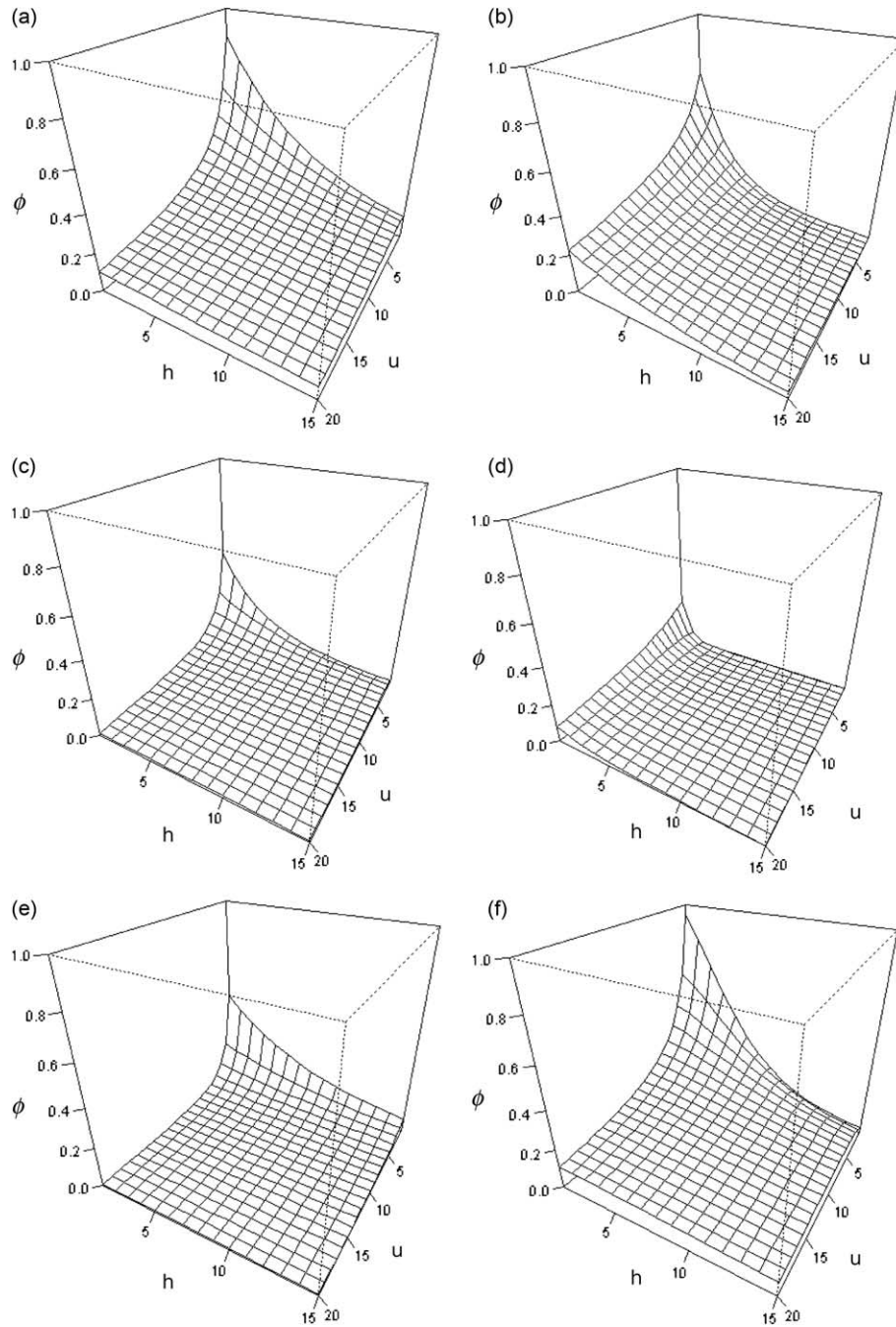


Figure 5. Theoretical space-time covariance functions, using reasonable values of key parameters. The first five panels (a–e) refer to  $\nu = 0.5$ , while in the sixth (f)  $\nu = 1.5$ . The default parameters are:  $\sigma^2 = 1$ ,  $d = 1.5$ ,  $c = 0.5$ ,  $\alpha = 1.2$  and  $\beta = 0.5$  (a), where in (b)  $\alpha = 0.8$ , (c)  $\beta = 0.9$ , (d)  $c = 2$  and in (e)  $d = 3$ .  $\phi$  is the axis of the covariance,  $h$  is the axis of spatial lags, and  $u$  is the axis of temporal lags.

gentle variation in space (Fig. 5b), while increasing the dimensional parameter,  $\beta$ , crushes the curve and reduces the range and slope in the time dimension (Fig. 5c).  $\beta$  has a similar effect to  $d$  (Fig. 5e) and for this reason these two parameters are usually collapsed together. Finally, increasing the scaling parameter of space,  $c$ , increases the smoothness of the surface (very low variation) and increases the slope at the origin along the space dimension (Fig. 5d). These four factors determine translation along the three axes but do not change the overall form. In the last example

(Fig. 5f) a value of  $\nu = 1.5$  was used: the curve becomes convex allowing for large variation in the space dimension along the time axis.

Equation (6) was fitted to the tsetse data using a maximum likelihood method and cross validation analysis (see below), giving a range of possible values for each parameter. The combination of values that maximised the likelihood between the experimental and theoretical covariance and minimised the cross-validation error of the estimate was selected.

## Space-time prediction

In the absence of any spatial trend, the model is reduced to a space-time ordinary kriging one (STOK). The estimated Apparent Density at site  $s$  and time  $t$  is given by:

$$\hat{AD}(s; t) = w(s; t) + \hat{R}(s; t)$$

$$\hat{R}_{STOK}(s; t) = \sum_r \sum_q \lambda_{r,q}(s; t) R(s_q; t_r) \quad (7)$$

where  $w(s; t)$  is the annual cycle extracted by wavelet decomposition (different for each trap site and composed of T values);  $R(s; t)$  are the residuals produced after de-seasonalising the data, assumed as an independent zero mean spatio-temporal process; and finally the weights  $\lambda_{r,q}$  sum to 1 and are calculated by solving a system of linear algebraic equations depending on the space-time covariance function. This constraint allows non-biased predictions at the sampled locations. The  $\hat{\phantom{x}}$  symbol means estimated value.

Removing seasonality reduces some of the variance and a large part of the distribution asymmetry (Table 3), but both are still high, indicating the presence of some other factor affecting the randomness of  $R$ .

The second model, RSTSK, is composed of a linear trend, modelled using  $\bar{W}$  and the fragmentation indices, and the residuals  $R_R$  modelled by spatio-temporal simple kriging.

The RSTSK linear regression is chosen after comparing different regression equations in relation to the dependent variable (the longitudinal  $AD$  for each trap or the temporal mean of  $AD$  for each trap,  $\bar{AD}$ ), and predictors: fragmentation indices alone or fragmentation indices plus  $\bar{W}$  or the trap mean of the wavelet coefficients ( $\bar{W}$ ). The first two regressions in Table 4 (rows 1, 2, 5 and 6) are the regressions considering the whole longitudinal data, while the last two are those that consider the temporal mean of  $AD$  and  $\bar{W}$  per trap.  $CA$ ,  $MPS$  and  $PSSD$  are static values and do not change in time (Table 4, rows 3, 4, 7 and 8). As indicated by the coefficient of determination ( $r^2$ ) the use of wavelet coefficients increases the efficiency of the linear regression (comparison between rows 1 and 2 versus 3, 4 and 5 and 6 versus 7, 8). The strongest correlation (indicated with a star in Table 4) is found using the average temporal Apparent Density as the dependent variable

and the average temporal wavelet coefficients, total area suitable habitat, mean patch size and patch size standard deviation as predictors.  $\bar{W}$  and  $CA$  presented the largest coefficients while  $MPS$  and  $PSSD$  contributed only a very small proportion to the overall prediction of  $AD$ .

In the second stage of RSTSK, using the mean of the linear regression,  $R_R$  for *G. p. gambiensis* (subscript pal) was estimated as follows:

$$R_{Rpal}(s; t) = AD_{pal}(s; t) - (16.48 + 19.68 \times \bar{W} - 2.67 \times CA + 0.26 \times MPS + 0.09 \times PSSD) \quad (8)$$

and for *G. tachinoides* (subscript tac):

$$R_{Rtac}(s; t) = AD_{tac}(s; t) - (-0.97 + 13.94 \times \bar{W} + 0.37 \times CA + 0.01 \times MPS - 0.01 \times PSSD) \quad (9)$$

In contrast to  $R$  derived through de-seasonalising (see above), the residual  $R_R$  derived through linear regression presents acceptable mean, variance, and asymmetry values close in shape to the standard Gaussian curve (Table 3).

The predicted  $\hat{AD}$  is then obtained summing the local spatial mean ( $\hat{m}$ ) and the interpolated residuals  $\hat{R}_R$ :

$$\hat{AD}(s; t) = \hat{m}(s) + \hat{R}_R(s; t)$$

$$\hat{R}_R(s; t) = \sum_r \sum_q \lambda_{r,q}(s; t) R_R(s_q; t_r) \quad (10)$$

where the weights  $\lambda_{r,q}$  are not constrained to sum to 1, because of the local variation of the mean. Apart from relaxing the kriging weights constraint, eq. (7) and (10) are very similar with the additional change that wavelet seasonality (explicit in eq. (7)) is now part of the local spatial mean in eq. (10).

## Validation and test

Two statistics were used to compare the two methods: the correlation coefficient between the predicted and the original data (De Iaco et al. 2005) and the D<sup>2</sup> validation criterion (Sahu and Mardia 2005). Three other statistics were applied in order to validate the kriging estimation by a leave-one-out cross validation method: the mean error (ME), the mean squared error (MSE) and the mean squared deviation ratio (MSDR) (Webster and Oliver 2007).

Table 4. Regression models to predict the monthly Apparent Density per trap (i.e. location) ( $AD$ ) or its temporal mean ( $\bar{AD}$ ) per trap/location of *G. palpalis gambiensis* and *G. tachinoides*. Y is the dependent variable for each trap  $s$ ,  $r^2$  the coefficient of determination,  $r^2_{adj}$  the adjusted  $r^2$  and SE the standard error of estimate. The co-regressors are:  $W$  wavelet coefficients per trap (or their mean through time  $\bar{W}$ ),  $CA$  total area of forest,  $MPS$  mean patch size and  $PSSD$  patch size standard deviation. The following equations are referred for *G. p. gambiensis* in Douroula and Kadomba, and *G. tachinoides* in Douroula.

Row	Y	Regression equations (Y ~ ...)	$r^2$	$r^2_{adj}$	SE
<i>G. p. gambiensis</i>					
1	$AD$	$12.97 + 0.32 \times W - 0.06 \times CA - 0.06 \times MPS - 0.06 \times PSSD$	0.10	0.09	17.65
2	$AD$	$12.97 - 0.06 \times CA - 0.06 \times MPS - 0.06 \times PSSD$	0.08	0.08	17.75
3	$\bar{AD}^*$	$16.48 + 19.68 \times W - 2.67 \times CA + 0.26 \times MPS + 0.09 \times PSSD$	0.71	0.70	5.35
4	$\bar{AD}$	$24.58 - 4.02 \times CA + 0.17 \times MPS + 0.15 \times PSSD$	0.48	0.43	9.65
<i>G. tachinoides</i>					
5	$AD$	$1.47 - 0.63 \times W + 0.01 \times CA + 0.01 \times MPS + 0.01 \times PSSD$	0.14	0.12	3.16
6	$AD$	$1.47 + 0.01 \times CA + 0.01 \times MPS + 0.01 \times PSSD$	0.05	0.04	3.31
7	$\bar{AD}^*$	$-0.97 + 13.94 \times W + 0.37 \times CA + 0.01 \times MPS - 0.01 \times PSSD$	0.68	0.67	0.43
8	$\bar{AD}$	$-1.55 + 0.54 \times CA + 0.08 \times MPS - 0.01 \times PSSD$	0.32	0.21	1.24



The statistics were calculated as follows:

$$ME = \frac{1}{NT} \sum_{s=1}^N \sum_{t=1}^T y(s; t) - \hat{y}(s; t) \quad (11)$$

$$MSE = \frac{1}{NT} \sum_{s=1}^N \sum_{t=1}^T (y(s; t) - \hat{y}(s; t))^2 \quad (12)$$

$$MSDR = \frac{1}{NT} \sum_{s=1}^N \sum_{t=1}^T \frac{(y(s; t) - \hat{y}(s; t))^2}{\sigma^2(s; t)} \quad (13)$$

$$D^2 = (AD - \hat{AD})^T \hat{\Sigma}^{-1} (AD - \hat{AD}) \quad (14)$$

with  $y=R$  in STOK,  $y=R_R$  in RSTSK,  $N$  is the total number of trap sites,  $\hat{\Sigma}$  is the covariance matrix of  $\hat{AD}$ , and  $\sigma^2$  is the space-time kriging variance.  $D^2$ , unlike the other measures, takes into account the spatial and temporal dependence between the observations.

## Results of space-time covariance analysis

As shown in Table 3, the  $R$  residuals are high asymmetric (high skew and kurtosis). Because the covariance function is sensitive to departures from normality, the values of  $R$  were log-transformed before further analysis (Tan et al. 2004). This was not necessary for the  $R_R$  values, which were already near normally distributed. The transformed residuals  $R$  and the non-transformed  $R_R$  were modelled, using eq. (6), as a function of a spatio-temporal random field process, with coefficients listed in Table 5, and then kriged.

The residuals  $R$  and  $R_R$  show a structured isotropic space-time covariance for 10 spatial lags of 200 m and 8 time lags of 1 month. As described above (space-time prediction section) these values are those that maximise the maximum likelihood between experimental and theoretical covariance function and minimise the kriging error. The space-time covariance is a decreasing function (opposite to the variogram, which is an increasing function) and the range within the covariance is both space and time dependent. The experimental covariance functions are isotropic because the spatio-temporal ranges and sills do not change according to the direction (not shown). First consider the residuals after spatial de-trending ( $R_R$ ). It is clear that the two tsetse species have the same spatial range (900 m) but different temporal ranges (longer in *G. p. gambiensis*). *Glossina tachinoides* seems to be less time dependent than is *G. p. gambiensis*, because of the shorter temporal range (5 months) and the absence of significant differences in its seasonality from site to site. The temporal range almost coincides with the length of the dry and

wet seasons. The shorter spatial range of  $R_R$  compared with  $R$  can be explained by the effect of removing the spatial trend in  $R_R$  (Cressie 1993).

## Accuracy, validation and prediction

The accuracy between the estimated and sampled  $AD$  was measured by  $D^2$  and by Pearson's coefficient of correlation ( $p$ ). The  $D^2$  statistic was tested approximating its distribution to a chi-square ( $\chi^2$ ) with  $N-1$  degrees of freedom by the Kolmogorov-Smirnov test. The  $p$ -value giving the probability that the result's distribution is a  $\chi^2$  is always over 0.50 with a maximum for *G. tachinoides* (0.89) using the RSTSK method. This indicates that RSTSK produced estimates that are closer than those of STOK to the target distribution and that there is no significant difference between the observed and estimated values. Summing the  $D^2$  values it is possible to produce a measure of the overall distance between predicted and observed data. *Glossina p. gambiensis* with RSTSK technique had the lowest value of  $D^2$ , indicating the best fit. The coefficient of correlation between the original Apparent Density and the estimated values is higher than 0.7 for almost all the variables analysed (only *G. tachinoides* with STOK technique shows a lower correlation,  $p=0.56$ ), confirming the robustness of combining techniques (wavelet analysis with ordinary space-time kriging in STOK and linear regression with space-time simple kriging in RSTSK).

Validation was performed by cross validation in space-time kriging, obtaining estimates for  $R$  (in STOK) or  $R_R$  (in RSTSK) (Table 6). ME and MSE should be close to 0 and MSDR to 1. Both methodologies resulted in satisfactory predictions for the two species; the optimal value for ME was reached in all cases and the MSDR was not far from unity, with better results in RSTSK. In general the algorithms are more efficient for *G. p. gambiensis* than for *G. tachinoides*, in part because of the larger sample.

The overall result shows that the RSTSK technique gave the best predictions. An example of the predictions given by the two frameworks is showed in Fig. 6 for May 2007 in Douroula. The RSTSK and STOK predictions for *G. tachinoides* are not very different. For *G. p. gambiensis* the RSTSK is more accurate than STOK in predicting the lowest values in the south of the area.

## Discussion and conclusion

The results described in this paper show that tsetse abundance as measured by the  $AD$  of trap catches varies both in space and through time and that these variations are

Table 5. Theoretical nonseparable space-time covariance parameters.  $h$ , the pure spatial range;  $u$ , the pure temporal range;  $c$ , a scaling parameter;  $\alpha$  a smoothing parameter for time and  $\beta$  a parameter determined by the dimensionality of the data.

<i>Glossina</i> ssp.	Variable	Method	$c$	$\alpha$	$\beta$	$h$	$u$
<i>p. gambiensis</i>	$R$	STOK	200	1.50	0.99	1400	6
	$R_R$	RSTSK	1	1.50	0.50	900	6
<i>tachinoides</i>	$R$	STOK	8	1.90	0.70	1200	5
	$R_R$	RSTSK	1	1.70	0.85	900	5

Table 6. Summary statistics from cross validation (ME, mean error; MSE, mean squared error; MSDR, mean squared deviation ratio) and Pearson correlation coefficient  $p$  (level of significance = 95%) between the raw data and the values estimated by the two techniques applied.

Species	STOK on R				RSTSK on $R_R$			
	$p$	ME	MSE	MSDR	$p$	ME	MSE	MSDR
<i>G. p. gambiensis</i>	0.71	-0.02	0.81	1.48	0.81	0.00	0.90	0.98
<i>G. tachinoides</i>	0.56	-0.54	1.73	1.75	0.75	0.03	1.05	1.01

correlated with each other in seasonally varying ways. This approach is novel in tsetse studies because it quantifies the spatial and temporal autocorrelations of standard tsetse AD data, an indispensable requirement for geostatistical modelling and predictions. Also, autocorrelation is one of the principal sources of bias in residuals and thus should be accounted for by spatial (and temporal) ecological models. Some of the results are individually very similar to those from previous work (for example the spatial range in Guerrini and Bouyer 2007), but this new procedure has the three-fold advantage of 1) exploring the simultaneous autocorrelation in space and time, 2) defining the spatial range and temporal range independently and then the relationship between them and 3) making predictions of fly abundance in un-sampled areas using a stochastic approach based upon these new analyses.

The knowledge of the spatial and temporal ranges can be used to define the optimal distance, in space and time, of sampling and control points (a choice that is usually made, and certainly constrained, by logistical and economical

considerations alone). In fact, recent work (Sciarretta et al. 2005) shows the practical use of autocorrelation in the optimal deployment of traps for tsetse control strategies, although the study referred to effectively considered the space and time dimensions independent of each other (traps were re-deployed depending only upon the current spatial distribution of tsetse, not on any changes over time); it is not, however, an integrated space-time autocorrelation analysis (Hohn et al. 1993).

The framework presented here separates and quantifies the different components influencing tsetse trap catches (seasonality, spatial trend, spatial and temporal autocorrelation ranges). The linear regression in RSTSK identified two important factors, the wavelet coefficients (seasonality) and the total area of forest as primarily determining changes in tsetse numbers through time and space. While the regression coefficients of wavelet values (themselves coefficients derived from wavelet analysis of the AD time series) are positive for both tsetse species, the equivalent coefficients for the total area of forest were positive for *G. tachinoides* (an expected result, given the literature on the importance of vegetation cover for tsetse; Bourn 1983, Hendrickx et al. 1999, Bouyer et al. 2005), but negative for *G. p. gambiensis*. These two species of tsetse, both in the *palpalis* group of flies, have rather different habitat requirements. Within the *palpalis* group, *G. tachinoides* is regarded as having the greater resistance to desiccation (equivalent to that of the savannah-dwelling *morsintans* group of flies) and might be expected to be more independent of habitat type than is *G. p. gambiensis*. This is also confirmed by the shorter temporal range, the stronger seasonality, and the lower correlation with fragmentation indices of *G. tachinoides* compared with *G. p. gambiensis*.

Thus the opposite effect of the total forest area on these two species remains enigmatic at present.

The present analysis allows predictions to be made of tsetse AD at times and places other than those of the actual samples taken, and within a reasonable spatial and temporal “window” around the current results (the autocorrelation cannot be generalized to other spatio-temporal geographic areas). One year of longitudinal data is not enough to create seasonal patterns for predictions over the longer term. A continuous period of at least three years of data collection should be regarded as the minimum to characterise the mean seasonal cycle and to obtain some measure on inter-annual variability (which is expected to be strong in the Sudan/Sahelian zone of the present study areas). The distribution of a variety of tsetse species appears to depend on subtleties of within – as well as between – annual variation in the climate cycle (Rogers and Robinson 2004) and future work is required to establish the links between these intra-annual cycles and the spatio-temporal analyses reported here.

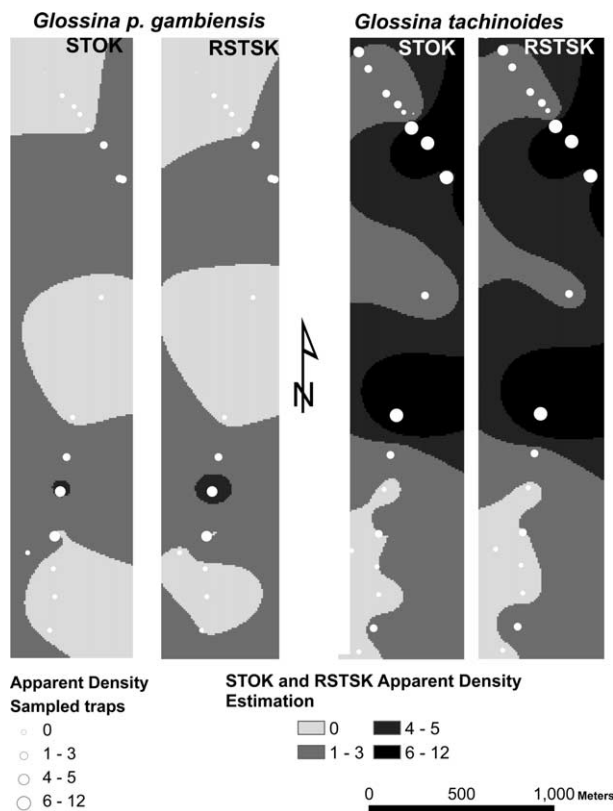


Figure 6. Sampled and estimated Apparent Density for Douroula in May 2007. The true values are shown for each trap point with graduated symbols (white points) on the STOK and RSTSK predictions respectively.

The model proposed is strictly statistical, based on the estimation of the implicit spatio-temporal autocorrelations for each species. It has shown that such spatio-temporal effects are important, but has not really explained why. Explanations now require the biologists to engage with the statisticians. Spatial and temporal trends in population density depend on environment conditions, food availability (Clausen et al. 1998, Torr and Mangwiro 2000) and on the biological processes of birth, death, immigration and emigration (Rogers et al. 1984) that all vary between the seasons. Furthermore, regulatory processes, such as density dependent birth and/or death, in theory occur irrespective of space, time or seasons but, in practice, are intricately associated with each. For example, tsetse numbers in such regions tend to peak at or near the end of the wet season; thus density dependence may also occur at this time of year only, and not at other times/seasons, when density is lower. Similarly, tsetse movement from one habitat to another may be in response to local habitat change, or to local tsetse abundance (Rogers et al. 1984).

A further analysis should consider biological processes such as density-dependent population change and movement in a non-restricted space-time model (i.e. non-stationary) for each fly sex separately (the Apparent Density used here merged male and female numbers). These can be included in the random field assumption, while the distribution and movements of cattle, sheep, goats etc. should be incorporated as additional co-variables in the linear regression of RSTSK, or as co-variables in the multi dimensional space-time covariance if such data are available.

The implications of these results for understanding better tsetse behaviour and ecology are obvious. Tsetse operate within spatial and temporal limits that can now be characterised fairly precisely by the sorts of space-time analyses carried out here. The implications for tsetse suppression and control are also important. Tsetse suppression often depends upon the expensive deployment of traps and or targets in an arrangement which is usually static (with trap spacing dependent upon guesswork combined with experience), but occasionally is adjusted adaptively in response to recent trap catches (Sciarretta et al. 2005). Reduction of trap deployment and maintenance costs during tsetse suppression programs has been a key aim because these costs significantly affect the overall cost of the operations. The present methods may be used to optimise initial trap placement and to decide whether or not such traps need to be moved seasonally. Whether or not the additional costs of working out trap re-deployment positions using the present methods, and of actually moving the traps, exceed the costs saved by so doing remains to be seen. Importantly the same techniques might be used to work out the limitations of trap improvements over present designs. A single trap, no matter how good it is, may not be able to attract and capture tsetse of the target populations over distances that are greater than the spatial ranges identified here; thus investment in making such a “super-trap” may be wasted by the diminishing returns of trapping out all the flies within the range of each trap.

Perhaps most importantly of all the present methods may be used to identify those parts of the habitat and times of year where and when tsetse concentrate. Targetting just

these areas for control may reduce tsetse numbers cost-effectively over much wider areas.

Current African Union plans to suppress or eradicate tsetse over large areas depend upon the concept of isolation of certain tsetse infested areas from each other. Whilst genetic techniques have been employed to identify the amount of genetic exchange (and hence effective mixing) that is occurring between two remote populations, the results are occasionally equivocal, and always depend upon a set of assumptions, some of which may not be met in the field. These genetic techniques could be complemented by the approach presented here. Whilst clearly we know what we know about tsetse behavior and ecology, at present we do not really know what we don't know that is important for such costly, large-area schemes to succeed in their endeavors to reduce the tsetse scourge across Africa. We hope the present technique helps to fill some of the gaps in our knowledge of these important species.

*Acknowledgements* – The authors thank J. P. Bord, E. Ducheyne, G. Hendrickx for providing information on habitat fragmentation, and H. Brown and D. Benz for reviewing the paper. Special acknowledgement is due to the Director General of CIRDES, Abdoulaye Gouro, for providing excellent working conditions, and to Félix Sanou, Bila Cene, Adama Sana, Sanogo Lansina and Joseph Sanou for their assistance during the field studies. Thanks are also due to T. Finnie, P. Gething, A. Patil, S. Randolph, and S. Serreli for useful discussions during the analysis of the data. This research was funded by the Wellcome Trust project 075824, “Fragfly”.

## References

- Arbogast, R. T. et al. 1998. Implications of spatial distribution of insect populations in storage ecosystems. – *Environ. Entomol.* 27: 202–216.
- Bogaert, P. 1996. Comparison of kriging techniques in a space-time context. – *Math. Geol.* 28: 73–90.
- Bourn, D. 1983. Tsetse control, agricultural expansion and environmental change in Nigeria. – Ph.D. thesis, Oxford Univ., UK.
- Bouyer, J. et al. 2005. A phyto-sociological analysis of the distribution of riverine tsetse flies in Burkina Faso. – *Med. Vet. Entomol.* 19: 372–378.
- Bouyer, J. et al. 2007a. A diffusion model of *Glossina p. gambiensis* (Diptera: Glossinidae) in Burkina Faso. – In: Vreysen, M. J. B. et al. (eds), Area-wide control of insect pests. From research to field implementation. Springer, pp. 221–228.
- Bouyer, J. et al. 2007b. Population structuring of *Glossina palpalis gambiensis* (Diptera: Glossinidae) according to landscape fragmentation in Mouhoun River, Burkina Faso. – *J. Med. Entomol.* 44: 788–795.
- Brightwell, R. et al. 1997. Changes over twelve years in populations of *Glossina pallidipes* and *Glossina longipennis* (Diptera: Glossinidae) subject to varying trapping pressure at Nguruman, south-west Kenya. – *Bull. Entomol. Res.* 87: 349–370.
- Cano, J. et al. 2007. Spatial and temporal variability of the *Glossina palpalis palpalis* population in the Mbini focus (Equatorial Guinea). – *Int. J. Health Geogr.* 6: 36.
- Christakos, G. 2000. Modern spatiotemporal geostatistics. – Oxford Univ. Press.

- Clausen, P. H. et al. 1998. Host preferences of tsetse (Diptera: Glossinidae) based on blood meal identifications. – *Med. Vet. Entomol* 12: 169–180.
- Cressie, N. A. C. 1993. Statistics for spatial data. – Wiley.
- Cressie, N. A. C. and Huang, H. 1999. Classes of nonseparable, spatiotemporal stationary covariance functions. – *J. Am. Stat. Assoc.* 94: 1330–1340.
- De Iaco, S. et al. 2005. Modeling and prediction of multivariate space-time random fields. – *Comput. Stat. Data Anal.* 48: 525–547.
- de la Rocque, S. et al. 2001. The changing distribution of two riverine tsetse flies over 15 years in an increasingly cultivated area of Burkina Faso. – *Bull. Entomol. Res.* 91: 157–166.
- Debnath, L. 2002. Wavelet transforms and their applications. – Birkhäuser.
- Diggle, P. J. 1990. Time series. – Oxford Univ. Press.
- Gneiting, T. 2002. Nonseparable, stationary covariance functions for space-time data. – *J. Am. Stat. Assoc.* 97: 590–600.
- Guerrini, L. and Bouyer, J. 2007. Mapping African animal trypanosomiasis risk: the landscape approach. – *Vet. Italiana* 43: 643–654.
- Guerrini, L. et al. 2008. Fragmentation analysis for prediction of suitable habitat for vectors: the example of riverine tsetse flies in Burkina Faso. – *J. Med. Entomol.* 45: 1180–1186.
- Hargrove, J. W. 2000. A theoretical study of the invasion of cleared areas by tsetse flies (Diptera: Glossinidae). – *Bull. Entomol. Res.* 90: 201–209.
- Hendrickx, G. et al. 1999. A systematic approach to area-wide tsetse distribution and abundance maps. – *Bull. Entomol. Res.* 89: 231–244.
- Hendrickx, G. et al. 2001. A contribution towards simplifying area-wide tsetse surveys using medium resolution meteorological satellite data. – *Bull. Entomol. Res.* 91: 333–346.
- Hohn, M. E. et al. 1993. Geostatistical model for forecasting spatial dynamics of defoliation caused by the gypsy moth (Lepidoptera: Lymantriidae). – *Environ. Entomol.* 22: 1066–1075.
- Huang, C. et al. 2006. Sensitivity of landscape metrics to classification scheme. – *Int. J. Remote Sens.* 27: 2927–2948.
- Hwang, C. and Turner, B. D. 2005. Spatial and temporal variability of necrophagous Diptera from urban to rural areas. – *Med. Vet. Entomol.* 19: 379–391.
- Jordan, A. M. 1986. Trypanosomiasis control and African rural development. – Longman.
- Kaufman, L. and Rousseeuw, P. J. 1990. Finding groups in data: an introduction to cluster analysis. – Wiley.
- Krafsur, E. S. 2009. Tsetse flies: genetics, evolution, and role as vectors. – *Infect. Genet. Evol.* 9: 124–141.
- Kuzoe, F. A. S. and Schofield, C. J. 2005. Strategic review of traps and targets for tsetse and African trypanosomiasis control. – UNICEF/UNDP/Worldbank/WHO, <<http://apps.who.int/tdr/svc/publications/tdr-research-publications/strategic-review-traps-targets-tsetse>>.
- Maudlin, I. 2006. African trypanosomiasis. – *Ann. Trop. Med. Parasitol.* 100: 679–701.
- Minasny, B. and McBratney, A. B. 2007. Spatial prediction of soil properties using EBLUP with the Matérn covariance function. – *Geoderma* 140: 324–336.
- Odit, M. et al. 2006. Using remote sensing and geographic information systems to identify villages at high risk for rhodesiense sleeping sickness in Uganda. – *Trans. R. Soc. Trop. Med. Hyg.* 100: 354–362.
- Odulaja, A. et al. 2001. Spatial and temporal distribution for tsetse fly trap catches at Nguruman, southwest Kenya. – *Bull. Entomol. Res.* 91: 213–220.
- Percival, D. B. and Walden, A. T. 2000. Wavelet methods for time series analysis. – Cambridge Univ. Press.
- R Development Core Team 2008. R: a language and environment for statistical computing. – R Foundation for Statistical Computing.
- Rawlings, P. et al. 1993. The distribution of tsetse flies *Glossina morsitans submorsitans* and *G. palpalis gambiensis* (Diptera: Glossinidae) in the Gambia and the application of survey results to tsetse and trypanosomiasis control. – *Bull. Entomol. Res.* 83: 625–632.
- Rogers, D. J. and Robinson, T. P. 2004. Tsetse distribution. – In: Maudlin, I. et al. (eds), *The trypanosomiasis*. CAB International, pp. 139–179.
- Rogers, D. J. et al. 1984. Local variation in the population dynamics of *Glossina palpalis palpalis* (Robineau-Desvoidy) (Diptera: Glossinidae). I. Natural population regulation. – *Bull. Entomol. Res.* 74: 403–423.
- Rogers, D. J. et al. 1996. Predicting the distribution of tsetse flies in West Africa using temporal Fourier processed meteorological satellite data. – *Ann. Trop. Med. Parasitol.* 90: 225–241.
- Sahu, S. K. and Mardia, K. V. 2005. A Bayesian kriged-kalman model for short-term forecasting of spatio-temporal processes. – *J. R. Stat. Soc. C* 54: 223–244.
- Schlather, M. 2001. Simulation of stationary and isotropic random fields. – *R-News* 1: 18–20.
- Sciarretta, A. et al. 2005. Development of an adaptive tsetse population management scheme for the Luke community, Ethiopia. – *J. Med. Entomol.* 42: 1006–1019.
- Tan, W. D. et al. 2004. Using normal quantile plot to select an appropriate transformation to achieve normality. – *Comput. Stat. Data Anal.* 45: 609–619.
- Torr, S. J. and Mangwiro, T. N. C. 2000. Interactions between cattle and biting flies: effects on the feeding rate of tsetse. – *Med. Vet. Entomol.* 14: 400–409.
- Van den Bossche, P. 2001. Some general aspects of the distribution and epidemiology of bovine trypanosomiasis in southern Africa. – *Int. J. Parasitol.* 31: 592–598.
- Wagner, H. H. and Fortin, M. J. 2005. Spatial analysis of landscapes: concepts and statistics. – *Ecology* 86: 1975–1987.
- Webster, R. and Oliver, M. A. 2007. Geostatistics for environmental scientists, 2nd ed. – Wiley.
- Whitcher, B. 2007. waveslim: basic wavelet routines for one-, two- and three-dimensional signal processing. – Inst. for Mathematics Applied to Geosciences, <[www.image.ucar.edu/~whitcher/](http://www.image.ucar.edu/~whitcher/)>, <[www.image.ucar.edu/~whitcher/book/](http://www.image.ucar.edu/~whitcher/book/)>.
- Yu, P. et al. 1995. Integration of temporal and spatial models for examining the epidemiology of African trypanosomiasis. – *Prev. Vet. Med.* 24: 83–95.